AD-A094 407 CALIFORNIA UNIV LOS ANGELES COGNITIVE SYSTEMS LAB F/G 5/10
ON THE VALUE OF SYNTHETIC JUDGMENTS. REVISION.(U)
DEC 80 M BURNS, J PEARL N00014-78-C-0372
UNCLASSIFIED UCLA-ENG-CSL-8032-REV NL

1 OF 1
AD
A094407

END
DATE
FILMED
2 81
DTIC

COGNITIVE SYSTEMS LABORATORY

ON THE VALUE OF SYNTHETIC JUDGMENTS

MICHAEL BURNS
JUDEA PEARL

SCHOOL OF ENGINEERING AND APPLIED SCIENCE • UNIVERSITY OF CALIFORNIA, LOS ANGELES

UCLA-ENG-CSL-8032
June 1980
Revision I, December 1980

ON THE VALUE OF SYNTHETIC JUDGMENTS

Michael Burns and Judea Pearl

Technical Report

Work performed at Cognitive Systems Laboratory
School of Engineering and Applied Science
University of California, Los Angeles
Professor Judea Pearl, Principal Investigator

This work was supported in part by the
Engineering Psychology Programs, Office of Naval Research
Contract N00014-78-C-0372, Work Unit Number NR 197-049

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

(6) ON THE VALUE OF SYNTHETIC JUDGMENTS,

5. TYPE OF REPORT & PERIOD COVERED

(9) Technical Report

6. PERFORMING ORG. REPORT NUMBER

(14) UCLA-ENG-CSL-8032-REV

7. AUTHOR(s)

(10) Michael Burns and Judea Pearl

8. CONTRACT OR GRANT NUMBER(s)

(15) N00014-78-C-0372

9. PERFORMING ORGANIZATION NAME AND ADDRESS

University of California, Los Angeles
School of Engineering and Applied Science
Los Angeles, California 90024

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

NR 197-049

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research
800 N. Quincy Street
Arlington, Virginia 22217

12. REPORT DATE

(11) June 1980, revised Dec 9 1980

13. NUMBER OF PAGES

21 pages

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

(12) 31

15. SECURITY CLASS. (of this report)

UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decomposition principle; causal schemata; judgment validity; probabilistic information processing.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Decision-support technologies are founded on the paradigm that direct judgments are less reliable and less valid than synthetic inferences produced from more "fragmentary" judgments. Moreover, certain types of fragments are normally assumed to be more valid than others. In particular, judgments about the likelihood of a certain state of affairs given a particular set of data (diagnostic inferences) are routinely fabricated from judgments about the likelihood of that data given various states of affairs (causal inferences).

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

411482

and not vice versa. This study was designed to test the benefits of causal synthesis schemes by comparing the validity of causal and diagnostic judgments against "ground-truth" standards.

The results demonstrate that the validity of causal and diagnostic inferences are strikingly similar; direct diagnostic estimates of conditional probabilities were found to be as accurate as their synthetic counterparts deduced from causal judgments. The reverse is equally true. Moreover, these accuracies were found to be roughly equal for each causal category tested. Thus, if the validity of judgments produced by a given mode of reasoning is a measure of whether it matches the format of human semantic memory, then neither one of the causal or diagnostic schema is a more universal or more natural format for encoding knowledge about common, everyday experiences.

These finding imply that one should approach the 'divide and conquer' ritual with caution; not every division leads to a conquest, even when the atoms are cast in causal phrasings. Dogmatic decompositions performed at the expense of conceptual simplicity may lead to inferences of lower quality than those of direct, unaided judgments.

On the Value of Synthetic Judgments

The objective of this study is to investigate empirically the conditions under which synthetic conclusions produced from 'fragmentary' judgments are more valid than direct, unaided inferences. We define 'validity' as the proximity between an assertion and the actual experience upon which it is based. The implicit assumption that synthetic conclusions are more valid than their direct counterparts is the basis for advocating the usefulness of all decision aiding technologies. Quoting Slovic, Fischhoff, and Lichtenstein (1977):

> Most of these decision aids rely on the principle of divide and conquer. This "decomposition" approach is a constructive response to the problem of cognitive overload. The decision aid fractionates the total problem into a series of structurally related parts, and the decision maker is asked to make subjective assessments for only the smallest components. Such assessments are presumably simpler and more manageable than assessing more global entities. Research showing that decomposition improves judgment has been reported by Armstrong, Denniston & Gordon (1975), Gettys et al. (1973), and by [Edwards, Phillips, Hays, and Goodman (1968)].
>
> Critics of the decomposition approach would argue that many of the aids require assessments of quantities the decision maker has never thought about, and that these apparently simple assessments may be psychologically more complex than the original decision. In some situations, people may really know what they want to do better than they know how to assess the inputs required for the decision aid (p. 17-18).

A closer look at decision aiding techniques reveals that the structuring procedures used fall into two categories: cascading and inversion. Cascading entails the chaining of a sequence of local judgments to produce the global inference. Inversion involves converting the direction of certain relations to a format more compatible with the decision maker's conceptualization of the environment. A typical example of cascading would involve inferring the consequence of a long sequence of actions. That is normally done by separately considering the effect of each individual action in the chain. Similarly, the aggregation of pieces of evidence in a multi-stage inferencing task would be an instance of cascading. For example, in the practice of decision analysis, the quality of actions are invariably inferred from judgments about the desirability of the actions' consequences cascaded by judgments about the likelihood of those consequences. Decision analysts never accept direct judgments about preferences on actions.

The most prevalent example of inversion is the insistance of decision analysts that information connecting evidential data with the hypothesis be cast in causal phrasings. Judgments about the likelihood of a certain hypothesis given a particular set of data (diagnostic inferences) are routinely fabricated from judgments about the likelihood of that data given various states of affairs (causal inferences), and not vice versa (Edwards et al., 1968; Howard, 1968; Raiffa, 1969; Tribus, 1969).

The experiments of Armstrong, Denniston, and Gordon (1975) and Gettys, Michel, and Steiger (1973) were directed toward verifying the benefit of cascading inferences. Armstrong et al. had subjects answer almanac-type questions; they tried to estimate some quantity (e.g., the number of pounds of tobacco processed in the U.S. in 1972) about which they had little or no *a priori* knowledge. Some subjects attempted to answer the overall 'global'

question, while others were instructed how to break the problem down into smaller subproblems. They found that decomposing the global problem into subproblems was helpful, especially on those problems where the subject knew practically nothing beforehand. Gettys et al. have tested the validity of likelihood estimates in the context of multi-staged hierarchical inference tasks. They compared posterior odds mentally assessed by subjects with the posterior odds calculated from Bayes' theorem and based on the actual histograms displayed to the subjects. Oddly enough, in their first experiment, which involved relations among height, scores, gender, and majors of students, direct mental assessments proved almost as accurate as those computed from the optimum model. Only in their second experiment involving a version of the urn problem did superiority of synthetic judgments surface. The results of the Armstrong et al. and Gettys et al. studies suggest that synthetic cascading inferences may indeed be a useful device in some instances. However, the question of how fine a division to employ should be approached with caution. Whereas the child who is learning arithmetic normally views multiplication as a sequence of additional operations, such a view may be detrimental to the more advanced student. Similarly, the pianist ought not to view his movements as being composed of individual muscular activations, but rather as a pattern of global entities such as scales, chords, arpegios, and the like. In the same vein, one may argue that as the decision maker becomes more familiar with the task environment he may achieve a state where unaided global inferences become more valid than their synthetic counterparts.

This paper focuses on the issue of causal/diagnostic inversion. The impetus for the hypothesis that causal judgments are more natural than diagnostic judgments may come from the fact that in statistical applications P(data|hypothesis) typically is obtained directly from a so-called statistical model, like the

assumption that a set of observations is normally distributed within given parameters (Edwards et al., 1968). This asymmetry also underlies the celebrated urn model. However, it is not at all clear whether any bias in favor of causal schema exists in cases where a parametric statistical model is not obvious and where both P(data|hypothesis) and P(hypothesis|data) are inferred by accessing semantic memory about everyday experiences.

Tversky and Kahneman (1977) indeed detected what they called 'causal biases' in decision making. They showed that subjects perceive causal information to have a greater impact than diagnostic information of equal informativeness. Further, if some information has both causal and diagnostic implications, then subjects' judgments are 'dominated' by the causal rather than the diagnostic relationship. Granted that causal reasoning is more emphasized in ordinary inference tasks, the question of the conditions under which the causal mode of reasoning will lead to more valid inferences still remains.

Aside from its psychological interest, this question has also acquired technological import. One application has already been alluded to, that of guiding the procedures used by decision analysts in eliciting likelihood estimates. The second application concerns organization of knowledge-based computer expert systems (Feigenbaum, 1977). In this latter application judgments from experts are encoded in the form of heuristic rules which are later combined to yield expert-like conclusions, explanations, and interpretations. The appropriate format for these fragmentary judgments is still subject to debate. Some knowledge-based systems (e.g., Shortliffe's MYCIN, 1976) insist on diagnostic inputs. Others (e.g., Ben-Bassat's MEDAS, 1980) require the more traditional causal judgments. The issue is whether experts, such as physicians, find it more comfortable to estimate the likelihood of diseases given a set of symptoms or to evaluate the likelihood that a given disease be accompanied by a certain

set of symptoms. Comfort aside, which form of input yields more valid therapeutical recommendations?

The experiment reported in this paper was designed to shed light on some of these issues. The problem of testing judgment validity, which has long been exacerbated by the lack of suitable criteria for measuring the quality of judgments about real-life experiences, was circumvented by 'creating' our own ground-truth data.

Method

Subjects

One hundred-seven undergraduate engineering students and 58 graduate students from various departments at UCLA participated in this study. The undergraduates, who were enrolled in one of two upper level undergraduate engineering classes, served in the experiment as part of an in-class lecture. The graduate students were recruited via advertisements posted around the campus and in the campus newspaper, and they were paid according to the accuracy of their judgments.

Materials

The undergraduates participated in the first phase of the study. Their task was to answer 24 yes/no questions concerning their activities and beliefs; the answers provided the data base (ground truth) for the estimation phase which followed. The questions were of two types: 'X questions' and 'Y questions', equal in number and randomly ordered on the questionnaire. Each X query questioned a condition, activity, or belief considered by the experimenters to be a causal agent for a condition, activity, or belief specified in one Y query. For example, since the color of a person's eyes is perceived to be influenced by that person's parents, not vice versa, X may represent the event of a mother having blue eyes and Y may denote the condition of her daughter having blue eyes. Four categories of causal relations were employed: (1) genetic causality, where a genetic condition specified by the X question serves as a cause of the condition designated by the Y question; (2) training causality, where the X condition provides training for the Y activity; (3) habit-forming causality, where the X condition serves as a habit-forming agent for the behavior specified by the Y condition, and finally; (4) self-interest causality, where the X question defines a particular self-interest that leads to the belief unveiled by the Y question. Table 1 shows the four causal categories

and the corresponding X and Y questions for each category.

Insert Table 1 about here

The data compiled from the undergraduates' responses served as the estimation targets in the second phase of the experiment. In this phase, the graduate students' task was to estimate the proportion of undergraduates responding in particular ways on the questionnaire. For a given X-Y relation, each graduate student was instructed to estimate either a causal triplet or a diagnostic triplet. When estimating the causal triplet, the estimator first considered $P(X)$ (e.g., the proportion of undergraduates who said their mother had blue eyes), then $P(Y|X)$ (e.g., the proportion of those undergraduates who said their mother had blue eyes and also said they themselves have blue eyes), and then $P(Y|\overline{X})$ (e.g., the proportion of those undergraduates who said their mother did not have blue eyes, but said they themselves have blue eyes). In assessing a diagnostic triplet, the student first estimated $P(Y)$ (e.g., the proportion of undergraduates who said they have blue eyes), then $P(X|Y)$ (e.g., the proportion of those undergraduates who said they have blue eyes, and also said their mother had blue eyes), then $P(X|\overline{Y})$ (e.g., the proportion of those undergraduates who said they do not have blue eyes, but said their mother had blue eyes). Note that the three components of each triplet represent statistically independent quantities and, moreover, that every component can be deduced from the three members of the opposing triplet via Bayes' theorem.

Procedure

The undergraduates answered the questionnaire during a regularly scheduled class meeting. The graduate students were assembled in groups ranging in size from 4 to 15 persons. Before they began the task, the graduate students were

told about the nature of the estimations they would be making, and about the 'pay scale' which was dependent on the proximity of their estimates to the actual proportions computed from the undergraduates' responses. Half of the graduate students estimated causal triplets for odd-numbered X-Y relations and diagnostic triplets for even-numbered relations. For the other half of the subjects, this pattern was reversed.

Each graduate student estimated one triplet for each of the 12 relations, thus making a total of 36 probability estimates. The estimators were given as much time as needed to contemplate the estimates required. Most of the students took between 20 and 30 minutes to complete the task.

Results and Discussion

The task of evaluating judgment validity requires a choice of a validity criterion. A variety of criteria has been proposed and utilized for measuring the degree of disparity between a given actual proportion P_a and an estimate P_e of that proportion (Pearl, 1977). We have examined both the quadratic error:

$$Q = (P_e - P_a)^2 \tag{1}$$

and the logarithmic error:

$$L = P_a \log P_a/P_e + (1-P_a) \log (1-P_a)/(1-P_e) .$$

Both gave rise to practically identical patterns, so this paper will present data based on the quadratic error only.

For each query we took $\overline{Q}$, the arithmetic mean of the quadratic errors across subjects, as a measure of inaccuracy of the corresponding estimate. These mean quadratic errors, along with the actual proportions and mean estimates, are shown in Table 2. These estimates are called <u>direct estimates</u> to distinguish them from <u>synthetic estimates</u>, which will be discussed later.

Insert Table 2 about here

Table 2 reflects a slight trend for the mean estimates to regress toward the .50 probability level in relation to the actual probability. That is, in 65% of the cases, the proportions were actually 'more extreme' (closer to .00 or 1.00) than their associated estimates. This effect is more apparent in Figure 1, which displays the relationship between the actual proportions (along the horizontal axis) and their associated estimates (along the vertical axis).

Insert Figure 1 about here

By and large, one cannot detect a marked difference in accuracy between causal estimates (i.e., $P_e(Y|X)$ and $P_e(Y|\overline{X})$) and their diagnostic counterparts (i.e., $P_e(X|Y)$ and $P_e(X|\overline{Y})$). In Figure 1, for example, where accuracy is reflected by proximity to the diagonal line, the two families of estimates appear equally dispersed. However, such a comparison is not entirely reliable. Since the values of the actual proportions $P_a(Y|X)$ are generally smaller than those of $P_a(X|Y)$, a direct comparison between their estimates may not reflect true differences in validity. An estimation error in the neighborhood of P = .50 is far less severe than an error of equal magnitude near the extremes (.00 and 1.00). On four of the X-Y relations the actual proportions $P_a(Y|X)$ and $P_a(X|Y)$ are fairly close to one another (within .15). In all four cases $P_e(Y|X)$ is at least slightly more accurate than $P_e(X|Y)$, lending some support to the hypothesis that causal reasoning leads to better inference-making than diagnostic reasoning. However, if the same procedure is employed with the

$P_e(Y|\bar{X})$ estimate (invoking causal reasoning) and the $P_e(X|\bar{Y})$ estimate (based on diagnostic reasoning), only three of the seven comparisons show an advantage for $P_e(Y|\bar{X})$. Since the difference between causal and diagnostic reasoning in these 11 comparisons is generally of small magnitude, there is not a noticeable advantage for the former, as had been anticipated.

Another way to circumvent the 'apples versus oranges' difficulty is to synthesize causal and diagnostic estimates that can be compared on equal ground. To do this we aggregated subjects' estimates by Bayes' theorem to calculate synthetic estimates according to the following equations:

$$P_s(X) = P_e(X|Y)\ P_e(Y) + P_e(X|\bar{Y})\ [1-P_e(Y)] \tag{2}$$

$$P_s(Y) = P_e(Y|X)\ P_e(X) + P_e(Y|\bar{X})\ [1-P_e(X)] \tag{3}$$

$$P_s(Y|X) = \frac{P_e(X|Y)\ P_e(Y)}{P_s(X)} \tag{4}$$

$$P_s(X|Y) = \frac{P_e(Y|X)\ P_e(X)}{P_s(Y)} \tag{5}$$

$$P_s(Y|\bar{X}) = \frac{[1-P_e(X|Y)]\ P_e(Y)}{[1-P_e(X|Y)]\ P_e(Y) + [1-P_e(X|\bar{Y})]\ [1-P_e(Y)]} \tag{6}$$

$$P_s(X|\bar{Y}) = \frac{[1-P_e(Y|X)]\ P_e(X)}{[1-P_e(Y|X)]\ P_e(X) + [1-P_e(Y|\bar{X})]\ [1-P_e(X)]} \tag{7}$$

Note that the synthetic estimates in (2), (4), and (6) should be regarded as diagnostic since they are deduced from diagnostic inputs. Similarly, the estimates constructed by formulas (3), (5), and (7) are causal.

Furthermore, the synthetic estimates are more reflective of the transformations employed by common Decision Analysis procedures. For example, formula (5)

represents the celebrated transformation from prior to posterior which was pioneered (posthumously) by Reverend Bayes in 1761 as a means to infer the "probability of causes". It has since become almost a ritual to assume that this transformation automatically produces more valid judgments than the direct estimate $P_e(X|Y)$.

Table 3 shows the mean quadratic error, $\bar{Q}$, for both the direct estimate and the synthetic estimate for each of the four conditional probabilities. The direct estimates for $P(Y|X)$ and $P(Y|\bar{X})$ involve causal reasoning and the direct estimates for $P(X|Y)$ and $P(X|\bar{Y})$ involve diagnostic reasoning, while this relationship reverses when the synthetic estimates are considered.

Insert Table 3 about here

Also shown is an indicator called the <u>normalized error difference</u> which gives a measure of the significance of the difference between the direct estimate and the synthetic estimate. It was computed by the following formula:

$$\text{normalized error difference} = \frac{\sqrt{n}(\bar{Q}_{diagnostic} - \bar{Q}_{causal})}{\sqrt{\sigma^2_{causal} + \sigma^2_{diagnostic}}} \tag{8}$$

where $\bar{Q}_{diagnostic}$ and $\bar{Q}_{causal}$ stand for the mean quadratic error across subjects for either the direct or synthetic estimates, as appropriate, and σ^2 represents the variance of those quadratic errors. One property of this normalized error difference is rather obvious: Its value is made increasingly positive when the validity of the causal estimate becomes significantly greater than that of the diagnostic estimate, and negative when the reverse is true. Clearly, since the same actual proportion applies to both the direct estimate and the synthetic estimate for a particular probability, the 'apples versus oranges' problem is eliminated.

Across all estimates, there are nine instances where the normalized error difference is significant at the .05 level according to a standard two-tailed t-distribution. In six of these cases, it is the causal estimate that is better than the diagnostic, which leaves three cases in which the diagnostic is better. Thus, there is little evidence in these data for the superiority of causal reasoning over diagnostic reasoning. In fact, only one of the problems (musical instrument) shows a positive normalized error difference for all four conditional probabilities, while one other problem (typing) has a negative normalized error difference for all four conditionals.

Aside from comparing causal and diagnostic estimates, Table 3 also enables us to compare the validities of direct versus synthetic estimates. A suspicion that the latter may be more valid than the former could be based on the argument that each synthetic estimate combines tne output of three knowledge sources. If these were independent mental processes in the sense that the estimator providing them would consult different data or invoke different procedures for their production, then one would be justified in hypothesizing superiority for synthetic estimates over their direct counterparts. Comparing the data, one finds that in five of the nine significant cases, synthetic estimates are better than their direct counterparts.

Table 4 shows the mean quadratic errors for direct estimates and synthetic estimates with questions grouped according to the type of causality implied in the X-Y relations. These were obtained by averaging the quadratic errors over the X-Y relations with each causal category. In general, genetic relations induce slightly more accurate estimates than training and habit-forming relations, while self-interest relations induce the worst estimates of all. This pattern is true for both direct estimates and synthetic estimates. For each causality category the synthetic estimates are more valid than the direct estimates of $P(X|Y)$, and less

valid for $P(Y|X)$. In each case, the more valid estimates are those based on causal reasoning, which lends support to the conjecture that causal reasoning is more naturally invoked in interpreting common observations. However, when considering the other four columns ($P(X)$, $P(Y)$, $P(Y|\bar{X})$, $P(X|\bar{Y})$), the pattern of results no longer reflects causal superiority.

Insert Table 4 about here

Conclusions

Admittedly, having ourselves adhered to the belief that causal reasoning is a more natural mode of inference-making, we were somewhat surprised that the results do not show a stronger validity differential in this direction. Taking Table 3, for example, the overall mean of the normalized error difference is equal to .25, which clearly does not support the hypothesis of general causal superiority. In the few X-Y relations where significant validity differentials were detected, there was not a sizable bias favoring the causal mode. Thus, if the validity of judgments produced by a given mode of reasoning is a measure of whether that mode matches the format of human semantic memory, then neither the causal nor diagnostic schema is a more universal or more natural format for encoding knowledge about common, everyday experiences. It appears that semantic memory contains both causal schema and diagnostic schema. Which is invoked for a particular observational relation may depend on the nature of the relation, the anticipated mode of usage, and the level of training or familiarity of the observer.

These findings imply that one should approach the 'divide and conquer' ritual with caution; not every division leads to a conquest, even when the resultant atoms are cast in causal phrasings. Forced transformations from

diagnostic to causal judgments performed at the expense of conceptual simplicity may lead to inferences of lower quality than direct, 'holistic' judgments.

References

Armstrong, J. S., Denniston, W. B., Jr., & Gordon, M. W. The use of the decomposition principle in making judgments. Organizational Behavior and Human Performance, 1975, 14, 257-263.

Ben-Bassat, M., et al. Pattern-based interactive diagnosis of multiple disorders: The MEDAS system. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1980, PAMI-2, 148-160.

Edwards, W., Phillips, L. D, Hayes, W. L., & Goodman, B. G. Probabilistic information processing systems: Design and evaluation. IEEE Transactions on Systems Science and Cybernetics, 1968, SSC-4, 248-265.

Feigenbaum, E. A. The art of artificial intelligence: I. Themes and case studies of knowledge engineering. Proc. 5th Int. Joint Conf. Artificial Intelligence, 1977, 2, 1014-1029.

Gettys, C., Michel, C., Steiger, J. H., Kelley, C. W., & Peterson, C. R. Multiple-stage probabilistic information processing. Organizational Behavior and Human Performance, 1973, 10, 374-387.

Howard, R. A. The foundations of decision analysis. IEEE Transactions on Systems Science and Cybernetics, 1968, SSC-4, 211-219.

Pearl, J. An economic basis for certain methods of evaluating probabilistic forecasts. International Journal of Man-Machine Studies, 1978, 10, 175-183.

Raiffa, H. Decision analysis. Redding, MA: Addison-Wesley, 1970.

Shortliffe, E. H. Computer-based medical consultations: MYCIN. New York: Elsevier, 1976.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Tribus, M. Rational descriptions, decisions, and designs. New York: Pergamon Press, 1969.

Tversky, A., & Kahneman, D. Causal schemata in judgments under uncertainty. In M. Fishbein (Ed.), Progress in social psychology. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Table 1

	X Questions	Y Questions
Genetic Causality		
1.	Mother has blue eyes.	Student has blue eyes.
2.	At least 1 of student's parents is left-handed.	Student is left-handed.
3.	Student is a male over 5'9".	Student played on high school basketball team.
Training Causality		
4.	Student took musical lessons as a child.	Student currently plays a musical instrument.
5.	Student ran or jogged regularly in high school.	Student currently runs or jogs regularly.
6.	Student took typing in high school.	Student types ≥40 words/min now.
Habit-forming Causality		
7.	Student attended church regularly in high school.	Student attends church regularly now.
8.	Student is currently married.	Student is wearing a wedding ring.
9.	Student's father was 'handy' around home.	Student changes his own oil in his car.
Self-Interest Causality		
10.	Student finds it financially difficult to complete his college studies.	Student favors UCLA increasing financial aid at expense of larger classes.
11.	Student's family finds medical expenses to constitute a substantial burden.	Student favors nationalized medical care plan.
12.	Student closely follows UCLA football.	Student favors UCLA building on-campus football stadium.

Table 2

Actual Proportions, Mean Estimates, and Mean Quadratic Errors for Direct Estimates

X-Y Relation	P(X)			P(Y)			P(Y|X)			P(X|Y)			$P(Y\|\overline{X})$			$P(X\|\overline{Y})$		
	Prop	Mean	$\overline{Q}$	Prop	Mean	$\overline{Q}$	Prop	Mean	$\overline{Q}$	Prop	Mean	$\overline{Q}$	Prop	Mean	$\overline{Q}$	Prop	Mean	$\overline{Q}$
1. Blue eyes	.20	.34	.053	.22	.25	.024	.76	.60	.067	.67	.49	.072	.09	.23	.048	.06	.26	.072
2. Left-handed	.15	.24	.037	.05	.17	.027	.19	.43	.115	.60	.46	.098	.02	.20	.064	.13	.17	.017
3. Basket-ball	.39	.38	.037	.08	.15	.041	.12	.26	.082	.56	.70	.091	.06	.11	.032	.38	.37	.042
4. Musical instr.	.56	.49	.054	.24	.34	.069	.33	.41	.062	.77	.66	.072	.13	.13	.009	.49	.34	.108
5. Running	.34	.30	.036	.20	.38	.094	.36	.74	.183	.62	.55	.070	.11	.29	.056	.27	.23	.034
6. Typing	.56	.39	.081	.23	.24	.031	.25	.48	.132	.60	.79	.066	.21	.22	.037	.55	.28	.115
7. Church	.36	.30	.041	.22	.28	.032	.49	.54	.056	.79	.78	.047	.07	.10	.013	.24	.34	.045
8. Wedding ring	.06	.16	.034	.04	.15	.040	.50	.68	.123	.75	.82	.079	.01	.02	.000	.03	.16	.048
9. Repairs	.74	.52	.097	.64	.37	.142	.66	.55	.086	.76	.62	.093	.57	.34	.107	.69	.47	.119
10. Finan. aid	.34	.48	.077	.38	.39	.067	.44	.62	.144	.39	.69	.148	.35	.28	.075	.30	.32	.070
11. Medical care	.26	.40	.106	.50	.57	.056	.57	.66	.072	.30	.68	.198	.48	.33	.085	.23	.24	.056
12. Foot-ball*	.46	.60	.046	.42	.59	.121	.57	.73	.057	.62	.81	.063	.29	.34	.063	.34	.18	.042

* Total of 25 subjects answered this question.

Table 3

Mean Quadratic Errors for Direct Versus Synthetic Estimates

	$P(Y\|X)$			$P(Y\|\bar{X})$			$P(X\|Y)$			$P(X\|\bar{Y})$		
X-Y Relation	Direct	Synth.	N.E.D.*	Direct	Synth.	N.E.D.*	Direct	Synth.	N.E.D.*	Direct	Synth.	N.E.D.*
1. Blue eyes	.067	.156	2.842	.048	.044	-.137	.072	.058	.553	.072	.027	1.537
2. Left-handed	.115	.060	-2.460	.064	.015	-3.954	.098	.093	.216	.017	.035	-1.453
3. Basket-ball	.082	.067	-.503	.032	.015	-.625	.091	.062	.968	.042	.042	0
4. Musical instr.	.062	.130	2.504	.009	.059	2.608	.072	.046	1.179	.108	.069	2.034
5. Running	.183	.140	-1.530	.056	.064	.403	.070	.053	.542	.034	.044	-.504
6. Typing	.132	.117	-.454	.037	.024	-.461	.066	.067	-.031	.115	.120	-.177
7. Church	.056	.060	.174	.013	.037	1.004	.047	.068	-.910	.045	.029	.669
8. Wedding ring	.123	.077	-2.227	.000	.014	1.538	.079	.036	2.059	.048	.023	2.747
9. Repairs	.086	.135	1.538	.107	.143	1.567	.093	.070	.782	.119	.131	-.522
10. Finan. aid	.144	.114	-1.071	.075	.044	-1.256	.148	.151	-.104	.070	.067	.122
11. Medical care	.072	.082	.255	.085	.075	-.317	.198	.159	1.140	.056	.061	-.159
12. Foot-ball **	.057	.125	2.393	.063	.092	.703	.063	.047	-.847	.042	.057	-.695

* Normalized Error Difference. ** Total of 25 subjects answered this question.

Table 4

Mean Quadratic Errors for Different Causal Categories

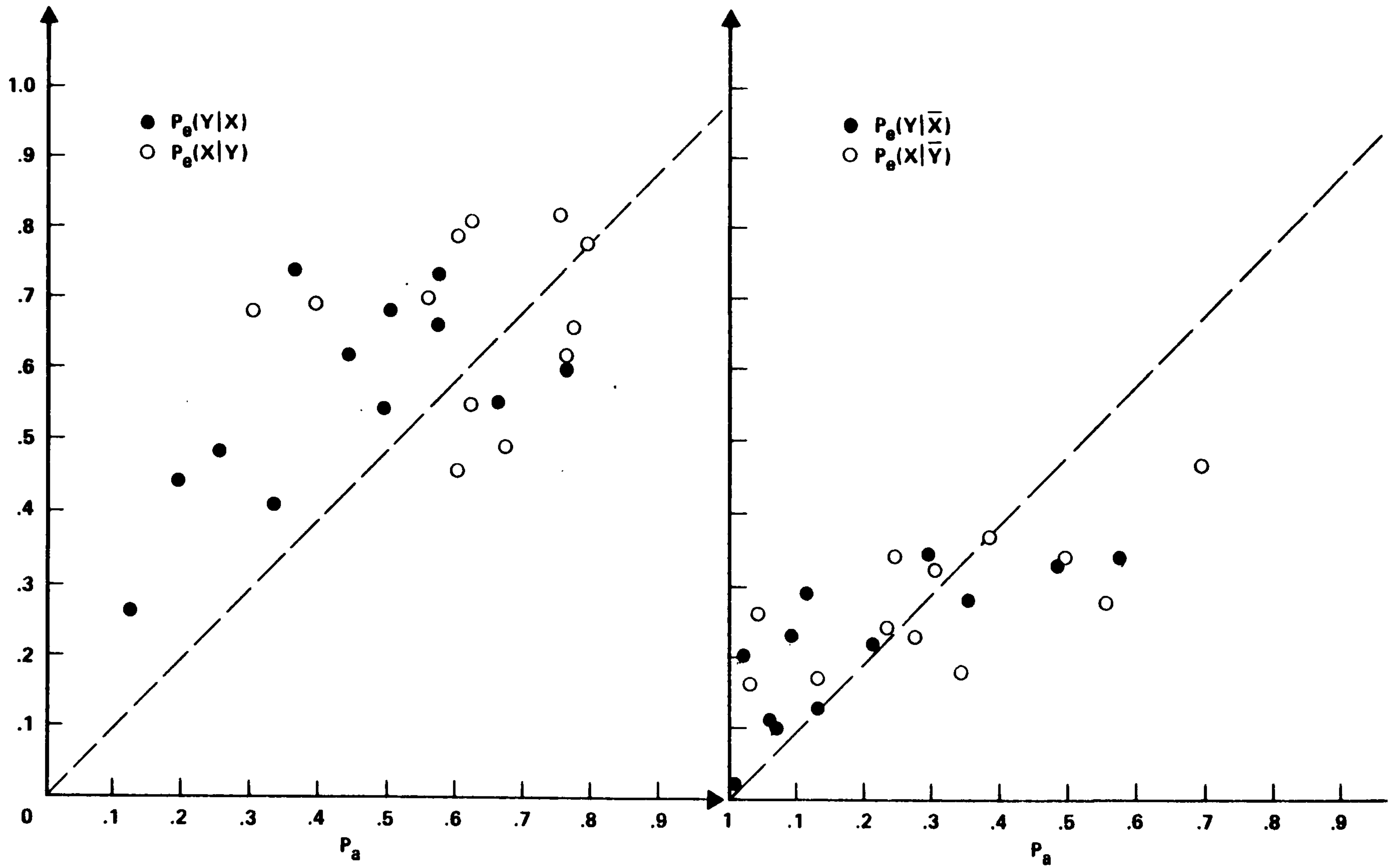
	$P(X)$		$P(Y)$		$P(Y\|X)$		$P(X\|Y)$		$P(Y\|\overline{X})$		$P(X\|\overline{Y})$		Mean	
	Dir.	Syn.	Dir.	Syn.	Dir.	Syn.	Dir.	Syn.	Dir.	Syn.	Dir.	Syn.	Dir.	Syn.
Genetic	.042	.038	.031	.056	.088	.094	.087	.071	.048	.025	.044	.035	.057	.053
Training	.057	.057	.065	.050	.126	.129	.069	.055	.034	.037	.086	.078	.073	.068
Habit-forming	.057	.073	.071	.043	.088	.091	.073	.058	.040	.060	.071	.061	.067	.064
Self-Interest	.076	.077	.081	.062	.091	.107	.136	.119	.074	.070	.051	.062	.086	.083
Mean	.058	.061	.062	.053	.098	.105	.091	.076	.049	.048	.065	.059		

Dir. = Direct Estimates

Syn. = Synthetic Estimates

Figure 1

Mean Estimates Versus Actual Proportions



DISTRIBUTION LIST

CDR Paul R. Chatelier
Office of the Deputy Under Secretary
of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Director
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217 (5 cys)

Director
Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Operations Research Programs
Code 434
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Statistics and Probability Program
Code 436
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Information Systems Program
Code 437
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Eastern/Central Regional Office
ATTN: Dr. J. Lester
Building 114, Section D
666 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. David
536 South Clark Street
Chicago, IL 60605

Commanding Officer
ONR Western Regional Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, CA 96503

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6 cys)

Dr. Robert G. Smith
Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Washington, D.C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation
Group
Naval Training Equipment Center
Code N-OOT
Orlando, FL 32813

Mr. Milon Essoglou
Naval Facilities Engineering Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

CDR Robert Biersner
Naval Medical R&D Command
Code 4
Naval Medical Center
Bethesda, MD 20014

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

CDR Thomas Berghage
Naval Health Research Center
San Diego, CA 92152

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Dr. James McGrath, Code 311
Naval Personnel Research and
Development Center
San Diego, CA 92152

Navy Personnel Research and
Development Center
Management Support Department
Code 210
San Diego, CA 92152

CDR P. M. Curran
Code 604
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Ronald A. Erickson
Human Factors Branch
Code 3194
Naval Weapons Center
China Lake, CA 93555

Human Factors Engineering Branch
Code 1226
Pacific Missile Test Center
Point Mugu, CA 93042

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, MD 21402

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. Warren Lewis
Human Engineering Branch
Code 8231
Naval Ocean Systems Center
San Diego, CA 92152

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 08D22
Washington, D.C. 20360

Commander
Naval Air Systems Command
Human Factors Programs
NAVAIR 340F
Washington, D.C. 20361

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, D.C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, D.C. 20360

Human Factor Engineering Branch
Naval Ship Research and Development
Center, Annapolis Division
Annapolis, MD 21402

LDCR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, CA 93940

Mr. Merlin Malehorn
Office of the Chief of Naval
Operations (OP 102)
Washington, D.C. 20350

Mr. J. Barber
HQS, Department of the Army
DAPE-MBR
Washington, D.C. 20310

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and
Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

U.S. Army Aeromedical Research Lab
ATTN: CPT Gerald P. Krueger
Ft. Rucker, AL 36362

ARI Field Unit-USAREUR
ATTN: Library
C/O ODCSPER
HQ USAREUR & 7th Army
APO New York 09403

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Air University Library
Maxwell Air Force Base, AL 36112

Dr. Gordon Eckstrand
AFHRL/ASM
Wright-Patterson AFB, OH 45433

North East London Polytechnic
The Charles Myers Library
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

Professor Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology
Establishment
Teddington, Middlesex TW11 OLN
ENGLAND

Director, Human Factors Wing
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND

Defense Technical Information Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 cys)

Dr. Craig Fields
Director, Cybernetics Technology
Office
Defense Advanced Research Projects
Agency
1400 Wilson Blvd.
Arlington, VA 22209

Dr. Judith Daly
Cybernetics Technology Office
Defense Advanced Research Projects
Agency
1400 Wilson Blvd.
Arlington, VA 22209

Professor Douglas E. Hunter
Defense Intelligence School
Washington, D.C. 20374

Dr. Robert R. Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, CA 93017

Dr. Gary McClelland
Institute of Behavioral Sciences
University of Colorado
Boulder, CO 80309

Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Dr. Miley Merkhofer
Stanford Research Institute
Decision Analysis Group
Menlo Park, CA 94025

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Professor Howard Raiffa
Graduate School of Business
Administration
Harvard University
Soldiers Field Road
Boston, MA 02163

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert Williges
Human Factors Laboratory
Virginia Polytechnical Institute
and State University
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, N.W.
Washington, D.C. 20036

Dr. Ward Edwards
Director, Social Science Research
Institute
University of Southern California
Los Angeles, CA 90007

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Kenneth Hammond
Institute of Behavioral Science
University of Colorado
Room 201
Boulder, CO 80309

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, N.W.
Washington, D.C. 20036 (3 cys)

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics
Company
St. Louis Division
Box 516
St. Louis, MO 63166

Mr. Richard J. Heuer, Jr.
27585 Via Sereno
Carmel, CA 93923

Dr. John Payne
Duke University
Graduate School of Business
Administration
Durham, NC 27706

Dr. Baruch Fischhoff
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. Andrew P. Sage
University of Virginia
School of Engineering and Applied
Science
Charlottesville, VA 22901

Dr. Lola Lopes
Department of Psychology
University of Wisconsin
Madison, WI 53706

DATE
FILMED

- 8